

Is Bell's theorem relevant to quantum mechanics? On locality and non-commuting observables

A. Matzkin

*Laboratoire de Spectrométrie Physique (CNRS Unité 5588), Université Joseph-Fourier Grenoble-1, BP 87,
38402 Saint-Martin d'Hères, France*

Abstract. Bell's theorem is a statement by which averages obtained from specific types of statistical distributions must conform to a family of inequalities. These models, in accordance with the EPR argument, provide for the simultaneous existence of quantum mechanically incompatible quantities. We first recall several contradictions arising between the assumption of a joint distribution for incompatible observables and the probability structure of quantum-mechanics, and conclude that Bell's theorem is not expected to be relevant to quantum phenomena described by non-commuting observables, irrespective of the issue of locality. Then, we try to disentangle the locality issue from the existence of joint distributions by introducing two models accounting for the EPR correlations but denying the existence of joint distributions. We will see that these models do not need to resort explicitly to non-locality: the first model relies on conservation laws for ensembles, and the second model on an equivalence class by which different configurations lead to the same physical predictions.

Keywords: Quantum mechanics, Non-locality, Bell inequalities

PACS: 03.65.Ta, 03.65.Ud, 02.50.-r

INTRODUCTION

The Bell inequalities result from Bell's theorem [1, 2]. This theorem is a mathematical statement, unrelated to any specific physical theory [3]. Briefly put, Bell's theorem in its simplest form tells us that average values obtained from a specific type of statistical distribution of a variable must conform to a family of inequalities. The specificity in question, coined under the questionable but widely used terminology "local hidden variables" (LHV), is to be found in the assumptions made in the derivation of the theorem. Its connection with quantum mechanics springs up from the dilemma put forward by Einstein, Podolsky and Rosen [4]: either (i) quantum mechanics is complete or (ii) physical quantities associated with non-commuting observables have simultaneous reality provided that locality holds. Indeed, LHV models adopt branch (ii) of the dilemma whose main assumption, is the 'simultaneous reality' of incompatible quantities – locality plays the role of an auxiliary assumption to avoid action at a distance. In Bell's theorem, this hazy terminology takes the form of a precise statement: the existence of a joint probability distribution for outcomes corresponding to incompatible observables [3, 5, 6, 7] – which in quantum mechanics only exists for commuting operators. In the first part of this note, we recall several well-known contradictions between alternative (ii) of the EPR dilemma and quantum mechanics. All these contradictions are grounded on the fact that incompatible physical quantities require a different probabilistic structure than the one offered by LHV models; in this sense, LHV models and the resulting Bell inequality are not relevant to quantum mechanics. The interesting question then is whether models complying with the quantum-mechanical requirement of non-commutativity are, in the context of EPR correlations, necessarily non-local. In the second part of this note, we will introduce two types of models which will turn out to be not necessarily non-local: in the first model, the EPR correlations can be attributed to a conservation law (in a holistic context however). The second model expels the locality issue from the EPR paradox by viewing a quantum state as an equivalence class of different but equivalent field/particle configurations.

BELL-TYPE MODELS AND INCOMPATIBLE PHYSICAL QUANTITIES

Bell's theorem

Bell's theorem (BT) is a mathematical statement giving a constraint on certain type of probability distributions, unrelated to any specific physical theory. We will nevertheless introduce the setting and the notation in line with

the two spin-1/2 particles in the singlet state system, which is the paradigmatic application of the Bell inequalities in quantum mechanics. We thus have two particles (formed by the fragmentation of an initial compound system) flying apart in opposite directions. A measurement, the spin projection along a chosen axis, is made on each of the particles. Let $i = 1, 2$ denote the particle, a, b the axis of the measurement (the "parameter" of the measurement) making respective angles θ_a, θ_b with an arbitrarily chosen z axis, and A_i, B_i the outcome obtained by measuring particle i along the axis a, b, \dots . Let us assume that particle 1 is measured along a and particle 2 along b . Each measurement can yield as possible outcomes $(A_1, B_2) = (\pm \frac{1}{2}, \pm \frac{1}{2})$ with observed frequencies $F(A_1, B_2)$. The resulting expectation value is

$$E(a, b) = \sum_{A_1, B_2} A_1 B_2 F(A_1, B_2) \quad (1)$$

where $A_1, B_2 = \pm \frac{1}{2}$.

Bell's theorem arises by supposing that each measurement is actually determined by an unknown variable λ that completely specifies the state of the system. Λ denotes the set containing all the λ 's, and $\rho(\lambda)$ the normalized distribution of the variable corresponding to a certain state of preparation of the system. Each λ gives rise to an outcome $(A_1(\lambda), B_2(\lambda))$ with a probability $p(A_1, B_2, \lambda)$. The observed frequencies are obtained by averaging over $\rho(\lambda)$

$$F_p(A_1, B_2) = \int p(A_1, B_2, \lambda) \rho(\lambda) d\lambda \quad (2)$$

and the expectation value $E(a, b)$ follows from

$$E_p(a, b) = \sum_{A_1, B_2} \int A_1 B_2 p(A_1, B_2, \lambda) \rho(\lambda) d\lambda. \quad (3)$$

To derive BT, one further assumption is needed, namely the factorisation of the joint probability $p(A_1, B_2, \lambda)$ in terms of two independent single particle probabilities,

$$p(A_1, B_2, \lambda) = p(A_1, \lambda) p(B_2, \lambda). \quad (4)$$

With this factorisation, condition, the expectation value takes the form

$$E_p(a, b) = \int \bar{A}_1(\lambda) \bar{B}_2(\lambda) \rho(\lambda) d\lambda, \quad (5)$$

where

$$\bar{A}_1(\lambda) = \sum_{A_1} A_1 p(A_1, \lambda) \text{ and } \bar{B}_2(\lambda) = \sum_{B_2} B_2 p(B_2, \lambda). \quad (6)$$

$\bar{A}_1(\lambda)$ (resp. $\bar{B}_2(\lambda)$) is the average over the outcomes A_1 (resp. B_2) obtained for a fixed value of λ . Indeed, in its most general form, λ does not determine the value of a given outcome A , but rather the probability $p(A, \lambda)$ of obtaining this outcome. This situation corresponds to *stochastic* Bell models. The so-called "deterministic" Bell models appear as a particular instance of the stochastic models when the probabilities $p(A_1, \lambda)$ and $p(B_2, \lambda)$ are all 0 or 1, in which case $\bar{A}_1(\lambda) = A_1(\lambda)$ and $\bar{B}_2(\lambda) = B_2(\lambda)$ meaning that a given λ univoquely determines the value of the measured outcomes. To obtain Bell's theorem, consider two directions a, a' for particle 1 measurements and two directions b, b' for particle 2 measurements (for simplicity all the directions are assumed to be coplanar). Then

$$|E(a, b) \mp E(a, b')| + |E(a', b) \pm E(a', b')| \leq 2V_{\max}^2, \quad (7)$$

where V_{\max} is the maximal value that can be taken by A or B (here, $\frac{1}{2}$). Eq. (7) is easily proven [1, 2] by making use of the factorization property (4) within each absolute value term $|\dots|$ and then employing triangle inequalities of the type $|\bar{B}_2 \mp \bar{B}_2'| + |\bar{B}_2 \pm \bar{B}_2'| \leq 2V_{\max}$.

Ruling out joint distributions

Let's forget about hidden variables for a moment to obtain two well-known inequalities. First, from the existence of a joint probability distribution (jd) $F(A_1, A'_1, B_2, B'_2)$, it is easy to recover the expectation values by marginalization, so that for example

$$E(a, b) = \sum_{A_1, B_2} A_1 B_2 \sum_{A'_1, B'_2} F(A_1, A'_1, B_2, B'_2). \quad (8)$$

Employing (8) and recalling that the absolute value of an average is bounded by the average of the absolute values, we have

$$|E(a, b) \mp E(a, b')| \leq \sum_{A_1 A'_1 B_2 B'_2} F(A_1, A'_1, B_2, B'_2) |A_1 (B_2 \mp B'_2)| \quad (9)$$

and the analog inequality for $|E(a', b) \pm E(a', b')|$. Adding both inequalities yields

$$|E(a, b) \mp E(a, b')| + |E(a', b) \pm E(a', b')| \leq \sum_{A_1 A'_1 B_2 B'_2} F(A_1, A'_1, B_2, B'_2) (|A_1 (B_2 \mp B'_2)| + |A'_1 (B_2 \pm B'_2)|) \leq 2V_{\max}^2, \quad (10)$$

where the right handside is obtained by using

$$|A_1 (B_2 \mp B'_2)| + |A'_1 (B_2 \pm B'_2)| \leq 2V_{\max}^2. \quad (11)$$

The second inequality is a quantum mechanical result valid for spin-1/2 projection operators. Let $\hat{S}_{1a}, \hat{S}_{1a'} \dots$ denote the operators whose eigenvalues correspond to the spin projections $A_1, A'_1 \dots = \pm V_{\max}$. A direct computation establishes that [8]

$$(\hat{S}_{1a}\hat{S}_{2b} \mp \hat{S}_{1a}\hat{S}_{2b'} + \hat{S}_{1a'}\hat{S}_{2b} \pm \hat{S}_{1a'}\hat{S}_{2b'})^2 = 4V_{\max}^4 \pm [\hat{S}_{1a}, \hat{S}_{1a'}][\hat{S}_{2b}, \hat{S}_{2b'}]. \quad (12)$$

This expression gives a bound for the norm of the operator between (...). Since $\|\hat{S}\| = V_{\max}$ the norm of each commutator is bounded by $2V_{\max}^2$, hence

$$\|\hat{S}_{1a}\hat{S}_{2b} \mp \hat{S}_{1a}\hat{S}_{2b'} + \hat{S}_{1a'}\hat{S}_{2b} \pm \hat{S}_{1a'}\hat{S}_{2b'}\| \leq 2\sqrt{2}V_{\max}^2, \quad (13)$$

and using the linearity of the operators and the fact that an expectation value (denoted $\langle \dots \rangle$, irrespective of the state) is bounded by the norm yields

$$|\langle \hat{S}_{1a}\hat{S}_{2b} \rangle \mp \langle \hat{S}_{1a}\hat{S}_{2b'} \rangle + \langle \hat{S}_{1a'}\hat{S}_{2b} \rangle \pm \langle \hat{S}_{1a'}\hat{S}_{2b'} \rangle| \leq 2\sqrt{2}V_{\max}^2. \quad (14)$$

Obviously if the commutators in Eq. (12) vanished, then Eq. (14) would be bounded by 2, just like the bound in BT. And in that case quantum mechanics allows to compute probabilities for joint events. But there is no joint distribution for non-commuting operators. Hence [3, 5, 6, 7], writing IRQ for 'irrelevant to quantum mechanics', one has

$$\textit{Bell's theorem is IRQ because it assumes joint distributions where quantum mechanics denies it.} \quad (15)$$

In the derivation of BT the factorization (4) is what brings in the existence of jd. Although (4) is known as Bell's locality condition, one can replace it with a *non-local factorizable* condition and still derive Bell's theorem [9]: non-local models also obey BT if they are factorizable.

Ruling out stochastic, then deterministic hidden variables

In the quantum context involving the fragmentation of two spin-1/2 particles formed in the singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|1+\rangle_u |2-\rangle_u - |1-\rangle_u |2+\rangle_u) \quad (16)$$

where u is any axis, the observed frequencies $F(A_1, B_2)$ are given according to quantum mechanics by the probabilities

$$P_\psi(A_1, B_2) = |\langle \psi | 1\text{sign}(A_1) \rangle_a |2\text{sign}(B_2) \rangle_b|^2. \quad (17)$$

In terms of the LHV, Eqs. (2) and (4) imply

$$F_p(A_1, B_2) = \int p(A_1, \lambda) p(B_2, \lambda) \rho(\lambda) d\lambda \quad (18)$$

where ρ is the distribution corresponding to the system having been prepared in the singlet state $|\psi\rangle$. Eq. (18) is inconsistent with stochastic Bell models. Indeed, choosing $b = a$ in Eq. (17) yields

$$P_\psi(A_1 = +\frac{1}{2}, B_2 \equiv A_2 = A_1 = +\frac{1}{2}) = 0, \quad (19)$$

$$P_\psi(A_1 = -\frac{1}{2}, B_2 \equiv A_2 = A_1 = -\frac{1}{2}) = 0, \quad (20)$$

while for the other 2 possibilities

$$P_\psi(A_1 = \pm\frac{1}{2}, A_2 = \mp\frac{1}{2}) = \frac{1}{2}. \quad (21)$$

Eqs. (19)-(21) yield the single particle probabilities $P_\psi(A_i) = 1/2$. Now summing Eq. (18) over A_1 or $B_2 = A_2$ gives

$$F_\rho(A_i = \pm\frac{1}{2}) = \int p(A_i = \pm\frac{1}{2}, \lambda) \rho(\lambda) d\lambda = \frac{1}{2} \quad (i = 1, 2). \quad (22)$$

But since $F_\rho(A_1, A_2 = -A_1)$ should also match (21), the expressions of the type $F_\rho(A_1) - F_\rho(A_1, A_2 = -A_1)$ vanish, from which it follows that

$$p(A_1 = \pm\frac{1}{2}, \lambda)(1 - p(A_2 = \mp\frac{1}{2}, \lambda)) = 0 \quad (23)$$

for any $\lambda \in \Lambda$, compatible only with unit or vanishing probability functions. Hence stochastic HV must be ruled out.

The only possibility is thus that of deterministic HV, but these must be ruled out as well. The argument goes back to Wigner [10], and is based on simple set theoretic assumptions – a set Λ over which a probability measure is defined is partitioned into different subsets having non-empty intersections [11]. A subset of Λ is in correspondence with an event, so that the measure of a subset represents the probability of the event. For example let Λ_{+a} denote the subset such that $A_1(\lambda) = +\frac{1}{2}$ (and hence $A_2(\lambda) = -\frac{1}{2}$; conversely for $\lambda \in \Lambda_{-a}$ we have $A_2(\lambda) = +\frac{1}{2}$ since the index is always relative to particle 1). These subsets cover the state space Λ such that

$$\Lambda = \Lambda_{+a} \cup \Lambda_{-a}. \quad (24)$$

Eq. (24) must be valid for any direction a so that if a' denotes an arbitrary axis, we have

$$\Lambda_{\pm a} = (\Lambda_{\pm a} \cap \Lambda_{+a'}) \cup (\Lambda_{\pm a} \cap \Lambda_{-a'}). \quad (25)$$

Assume now that B_2 has been measured and the outcome is known, say $B_2 = -\frac{1}{2}$. The quantum mechanical probabilities,

$$P_\psi(A_1 = \pm\frac{1}{2}, B_2 = -\frac{1}{2}) = \begin{cases} \frac{1}{2} \cos^2 \frac{\theta_b - \theta_a}{2} & \text{if } A_1 = +\frac{1}{2} \\ \frac{1}{2} \sin^2 \frac{\theta_b - \theta_a}{2} & \text{if } A_1 = -\frac{1}{2} \end{cases}, \quad (26)$$

should match the frequency predicted by LHV

$$F_\rho(A_1 = \pm\frac{1}{2}, B_2 = -\frac{1}{2}) = \int_{\Lambda_{+b}} p(A_1, \lambda) \rho(\lambda) d\lambda, \quad (27)$$

Since $p(A_1, \lambda)$ is 1 or 0 depending on whether $\lambda \in \Lambda_{\pm a}$, Eq. (49) becomes

$$F_\rho(A_1 = \pm\frac{1}{2}, B_2 = -\frac{1}{2}) = \int_{\Lambda_{+b} \cap \Lambda_{\pm a}} \rho(\lambda) d\lambda \equiv \mathcal{M}_{\Lambda_{+b} \cap \Lambda_{\pm a}}, \quad (28)$$

where $\mathcal{M}_{\Lambda_{+b} \cap \Lambda_{\pm a}}$ defines the measure of the subset $\Lambda_{+b} \cap \Lambda_{\pm a} \subset \Lambda$ (actually one can show that the rotational symmetry of the singlet state imposes that $\rho(\lambda)$ must be uniform and \mathcal{M} is simply the relative volume of $\Lambda_{+b} \cap \Lambda_{\pm a}$). If we measure particle 1's spin along a' , rather than along a , $F_\rho(A'_1 = \pm\frac{1}{2}, B_2 = -\frac{1}{2}) = \mathcal{M}_{\Lambda_{+b} \cap \Lambda_{\pm a'}}$. Using $\Lambda = \Lambda_{+a'} \cup \Lambda_{-a'}$ [Eq. (24)], we note that

$$\Lambda_{+b} \cap \Lambda_{+a} = (\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{+a'}) \cup (\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{-a'}) \quad (29)$$

so that

$$\mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{+a'}} = \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b}} - \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{-a'}}. \quad (30)$$

Using the trivial inequalities $\mathcal{M}_{\Lambda_{+a'} \cap \Lambda_{+b} \cap \Lambda_{-a'}} \leq \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{-a'}}$ and $\mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{+a'}} \leq \mathcal{M}_{\Lambda_{+a'} \cap \Lambda_{+b}}$ we infer from Eq. (30) that

$$\mathcal{M}_{\Lambda_{+a'} \cap \Lambda_{+b}} \geq \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{+a'}} \geq \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{+b}} - \mathcal{M}_{\Lambda_{+a} \cap \Lambda_{-a'}} \quad (31)$$

which, following (28) is an inequality corresponding to the probabilities predicted by the deterministic HV. However, this inequality is inconsistent with the quantum mechanical probabilities P_ψ : indeed according to Eq. (26), plugging in the P_ψ in Eq. (31) would lead to

$$\cos^2 \frac{\theta_b - \theta_{a'}}{2} \geq \cos^2 \frac{\theta_b - \theta_a}{2} - \sin^2 \frac{\theta_a - \theta_{a'}}{2}, \quad (32)$$

a relation that is not valid in general (eg it doesn't hold if we choose coplanar angles obeying $0 \leq \theta_b < \theta_a < \theta_{a'} \leq \pi/2$). Therefore, assuming deterministic HV leads to a contradiction, so that they must be ruled out as well; therefore

Bell's theorem is IRQ because neither stochastic nor deterministic factorizable LHV are consistent with quantum probabilities. (33)

This inconsistency is grounded on the quantity $\Lambda_{+a} \cap \Lambda_{+b} \cap \Lambda_{+a'}$, which does not correspond to any quantum-mechanical probability or associated quantity but is meaningful within the Bell-type deterministic models: this is the support for the joint events mentioned above, so that (33) appears as a consequence of (15).

LOCALITY AND NON-BELL-TYPE MODELS

The conclusion to be drawn from the preceding Section is that factorizability, implying the existence of joint distributions, is the origin of the inadequacy of Bell-type LHV models to account for quantum probabilities and expectation values. Non-local models can be factorizable (in which case they are also constrained by BT) or not. But what about *local* models? Since it is usually stated that Eq. (4) is the consequence of locality, it would appear that non-factorizable models cannot be local¹. In this section we challenge this assertion by giving an overview of two different types of models. The first model is built from the remark that the purported non-locality actually arises by the combination of non-commutative observables (precluding factorizability) and a conservation law (imposed by rotational invariance), so giving priority to non-locality or asserting that a conservation law is all that is needed becomes a matter of taste. The second model, based on particle and field configurations, defuses the EPR dilemma from the start: this model negates the existence of an element of reality from the possibility of making a prediction with unit-probability. As a consequence the model does not allow to efficiently complete quantum mechanics by a fully deterministic model ascribing sub-quantum probabilities.

Model 1: Conservation laws, holism or non-locality?

The model described in details elsewhere (see [13] and in particular Sec. IV of [14]) is based on ensemble properties of classical angular momenta distributions. Consider the fragmentation of an initial particle with a total angular momentum $\mathbf{J}_T = 0$ into 2 particles carrying angular momenta \mathbf{J}_1 and \mathbf{J}_2 . Conservation of the total angular momentum imposes $J_1 = J_2 \equiv J$ and

$$\mathbf{J}_1 + \mathbf{J}_2 = 0. \quad (34)$$

Without further constraints (or additional knowledge), the classical distribution in the 2-particle phase space is given by

$$\rho(\Omega_1, \Omega_2) = N \delta(\mathbf{J}_1 + \mathbf{J}_2) \delta(J_1^2 - J^2), \quad (35)$$

where N is a normalization constant. The corresponding distributions of the angular momenta in physical space – easier to visualize than ρ – is uniform on the angular momentum sphere, with \mathbf{J}_1 and \mathbf{J}_2 pointing in opposite directions. We can take $J = 1$ without loss of generality [13]. The detectors contain a random interaction and only deliver the results

¹ Bell was actually more precise – he carefully argued that Eq. (4) could be derived by assuming *local causality*, that is physical theories in which the measurement outcomes can be fully specified in terms of a complete set of beables [12]. This is stronger than requiring simple locality, which only involves the absence of action at a distance and not the issues of completeness and determinism.

$\pm \frac{1}{2}$. Let us take a closer look at the measurement process for a single particle whose \mathbf{J} distribution is ρ_{a+} , a uniform distribution on the hemisphere characterized by $J_a > 0$. Let $R_b = \pm \frac{1}{2}$ denote the outcomes for measurements along b . The system-apparatus interaction is assumed to verify the following property: the average over the outcomes R_b is equal to the mean value of the projection J_b over the initial distribution:

$$\langle R_b \rangle_{\rho_{a+}} = \sum_k k P(R_b = k, \rho_{a+}) = \langle J_b \rangle_{\rho_{a+}} = \frac{1}{2} \cos(\theta_b - \theta_a). \quad (36)$$

Three interesting properties follow. (i) Eq. (36) along with normalization is sufficient to impose the probabilities $P(R_b = \pm \frac{1}{2}, \rho_{a+}) = (\cos(\theta_b - \theta_a) \pm 1)/2$. (ii) Eq. (36) is inconsistent with the existence of elementary probabilities depending on \mathbf{J} , ie there can be no $p(R_b = \pm \frac{1}{2}, \mathbf{J})$ such that

$$P(R_b = \pm \frac{1}{2}, \rho_{a+}) = \int p(R_b = \pm \frac{1}{2}, \mathbf{J}) \rho_{a+}(\mathbf{J}) d\mathbf{J}, \quad (37)$$

as only ensemble-dependent elementary probabilities $p(R_b = \pm \frac{1}{2}, \mathbf{J}, \rho_{a+})$ are consistent with Eq. (36) [13, 14]. (iii) putting $b = a$ in Eq. (36) gives $P(R_a = \pm \frac{1}{2}, \rho_{a+}) = 1$ or 0; taking into account the ensemble dependency, this means that $R_a = 1/2 \Leftrightarrow J_a > 0$ for every $J_a \in \rho_{a+}$: when the distribution and measurement axes coincide, there is no interaction and the measurement device senses at most one hemisphere. This model is compatible eg with a particle following a stochastic motion with its angular momentum constrained to remain in the ensemble, the timescale of the measurement being significantly larger than the timescale of the stochastic motion.

We now return to the 2-particle problem with the uniform distribution ρ . Eq. (36) becomes $\langle R_{ia} \rangle_\rho = \langle J_{ia} \rangle_\rho = 0$ where $i = 1, 2$ and a is any axis. Eq. (34) and point (iii) above imply that the outcomes and the distributions for the particles along the same axis must be anti-correlated along any axis a

$$\langle J_{2a} \rangle_{\rho_{a\mp}} \equiv R_{2a} = -R_{1a} \equiv -\langle J_{1a} \rangle_{\rho_{a\pm}} \quad (38)$$

where the efficient distribution $\rho_{a\pm}$ in which the particle undergoes its stochastic motion depends on the initial position of the angular momentum and on the choice of the measurement axis. Measuring R_{1a} links the outcome to one of the two ensembles $\rho_{1a\pm}$ depending on whether $R_{1a} = \pm 1/2$. Note that contrarily to the correlation between individual phase-space positions (for which one has $J_{2a} = -J_{1a}$ and $J_{2b} = -J_{1b}$ jointly for any axes a and b), Eq. (38) cannot hold jointly along several directions (this is a consequence of the ensemble dependency, implying non-commutativity even for a single particle). Since the measurement outcomes do not depend on the individual phase-space positions, the average $E(a, b) \equiv \langle R_{1a} R_{2b} \rangle_\rho$ cannot be computed from phase-space averages, but from the probabilities of detecting a given outcome as a function of the distribution. $E(a, b)$ is computed from the general formula

$$\langle R_{1a} R_{2b} \rangle_\rho = \sum_{k, k' = -1/2}^{1/2} k k' P_{kk'} \text{ with } P_{kk'} = P(R_{1a} = k \cap R_{2b} = k', \rho) = P(R_{1a} = k) P(R_{2b} = k' | R_{1a} = k). \quad (39)$$

The two particle expectation takes the form

$$\langle R_{1a} R_{2b} \rangle_\rho = \sum_{k = -1/2}^{1/2} k P(R_{1a} = k) \left[\sum_{k' = -1/2}^{1/2} k' P(R_{2b} = k' | R_{1a} = k) \right]. \quad (40)$$

For any particle i and direction a , we have $P(R_{ia} = \pm \frac{1}{2}, \rho) = \frac{1}{2}$. The conditional probability $P(R_{2b} = k' | R_{1a} = k)$ is the probability of obtaining $R_{2b} = k'$ if it is known that $R_{1a} = k$. But obtaining an outcome $R_{1a} = k$ means that the distributions for particles 1 and 2 can be restricted to $\rho_{1a[\text{sign}(k)]}$ and $\rho_{2a[\text{sign}(-k)]}$ respectively. The conditional probability is therefore given by

$$P(R_{2b} = k' | R_{1a} = k) = P(R_{2b} = k', \rho_{2a[\text{sign}(-k)]}), \quad (41)$$

which is a single particle probability of the type given in point (i) below Eq. (36). Plugging these quantities into (40) leads to $E(a, b) = -\frac{1}{4} \cos(\theta_b - \theta_a)$, the quantum mechanical result for the singlet spin state (16).

The present model therefore does not abide by BT. The reason is twofold. First comes the ensemble dependency, enforcing not only non-commutativity, but the impossibility of ascribing elementary probabilities. Second comes the conservation of the angular momentum: what Eq. (38) does is to turn the conservation of the angular momentum over

the ensembles into the conservation of the angular momentum *between* these ensembles. This means that somehow, the particles must know what ensemble was picked by the first measurement in order to conserve the ensemble angular momentum of the second ensemble previous to its measurement. What is really happening is the application of the conservation law in the context of non-commutative measurements: contrarily to the commutative case where $J_{1a} = -J_{2a}$ and $J_{1b} = -J_{2b}$ can hold jointly, here $R_{1a} = -R_{2a}$ and $R_{1b} = -R_{2b}$ do not. At this point it would be possible to invoke non-locality to explain how the angular momentum can be conserved, though one can also uphold that conservation laws and symmetry principles are just postulated, without the need to invoke a specific mechanism. Alternatively it can be argued that symmetries can give rise to nonlocality, a position leading to a holistic vision of symmetries as holding beyond a space-time framework. Note that mechanical holistic systems – that is two systems maintaining a mechanical link between them – were already known to violate the Bell inequalities (an ad-hoc model was proposed in Ref. [15]). Here we have given a physical model that turns out to be the classical counterpart of quantum mechanical coupled angular momenta [14]: the violation of the Bell inequalities is necessary in order to conserve symmetries.

Model 2: Quantum states as equivalence classes

The model [16] represents a single spin-1/2 by a field-particle system composed of a small sphere, with the position of its center in the laboratory frame being denoted by \mathbf{x} and the internal spherical variables relative to the center of the sphere by $\mathbf{r} \equiv (r, \theta, \phi)$. A classical scalar field $F(\mathbf{r})$ is defined on the sphere's surface, while the point-like particle sits still at a fixed (but unknown position) on the sphere. As in Sec. 2, let B denote the spin projection along an axis b making an angle θ_b with the z axis. The outcome may depend on the position occupied by the field on the spherical surface and (ii) on the position of the particle. The field F is defined on the hemispherical surface centered on a given axis, the value of the field at any point being given by the projection of that point on the axis. Let Σ_{+a} denote the positive half-sphere centered on the axis a making an angle θ_a with the z axis, and $F_{\Sigma_{+a}}$ denote the field distributed on that hemisphere. $F_{\Sigma_{+a}}(\mathbf{r})$ is thus defined by

$$F_{\Sigma_{+a}}(\mathbf{r}) = \begin{cases} \mathbf{r} \cdot \mathbf{a} / \pi R^2 & \text{if } \mathbf{r} \in \Sigma_{+a} \\ 0 & \text{otherwise} \end{cases}, \quad (42)$$

R being the radius of the sphere (for simplicity we will take all the axes to be coplanar with z). The mean value of $\mathbf{r} \cdot \mathbf{b} / \pi R^2$ taken over Σ_{+a} is given by

$$\langle F_{\Sigma_{+b}} + F_{\Sigma_{-b}} \rangle_{\Sigma_{+a}} \equiv \int_{\Sigma_{+a}} \frac{\mathbf{r} \cdot \mathbf{b}}{\pi R^2} d\hat{\mathbf{r}} = \cos(\theta_b - \theta_a), \quad (43)$$

where $d\hat{\mathbf{r}}$ denotes the spherical surface element for a sphere of radius R . The only requirement we make on the particle's position is that it must be embedded within the field: the particle cannot be in a field free region of the sphere.

When a measurement is made we assume that the apparatus along b interacts with the field $F_{\Sigma_{+a}}$. Let $[a+b]$ and $[a-b]$ denote the directions lying halfway between the axes a (of the distribution) and b or $-b$ (of the measuring direction), with respective angles $(\theta_b + \theta_a)/2$ and $(\theta_b + \pi + \theta_a)/2$. We will assume that the field-apparatus interaction results in a *rotation* of the original pre-measurement field $F_{\Sigma_{+a}}$ toward both of the apparatus axes, $F_{\Sigma_{+a}} \rightarrow F_{\Sigma_{+b}} + F_{\Sigma_{-b}}$. A definite outcome $B = \pm \frac{1}{2}$ depends on which of the hemispheres $\Sigma_{\pm b}$ the particle is after the interaction. In terms of the field, this probability is given by the relative value of the average of the rotated field $F_{\Sigma_{+b}} + F_{\Sigma_{-b}}$ over the intermediate 'half-rotated' hemisphere $F_{\Sigma_{[a\pm b]}}$ depending on the initial field $F_{\Sigma_{+a}}$, yielding in accordance with Eq. (43)

$$P_{\Sigma_{+a}}(B = +\frac{1}{2}) = \left| \langle F_{\Sigma_{+b}} + F_{\Sigma_{-b}} \rangle_{\Sigma_{[a+b]}} \right|^2 / N = \cos^2 \frac{\theta_b - \theta_a}{2} \quad (44)$$

$$P_{\Sigma_{+a}}(B = -\frac{1}{2}) = \left| \langle F_{\Sigma_{+b}} + F_{\Sigma_{-b}} \rangle_{\Sigma_{[a-b]}} \right|^2 / N = \sin^2 \frac{\theta_a - \theta_b}{2} \quad (45)$$

with N being the sum of both terms. If b and a are taken to be the same, then one has $\Sigma_{[a+a]} \equiv \Sigma_{+a}$ and $P_{\Sigma_{+a}}(A = \pm \frac{1}{2}) = 1$ and 0 respectively. Hence a field $F_{\Sigma_{+a}}$ corresponds to a well-defined positive spin projection along the a axis. In this case the symmetry axis of the field distribution coincides with the post-measurement axis and the field-apparatus interaction may change the position of the particle though it remains within the hemisphere Σ_{+a} . On the

other hand when b and a lie along different directions, the spin projection along b only acquires a value $B = \pm \frac{1}{2}$ after the field has interacted with the measurement apparatus and rotated toward the measurement axis: the measurements do not commute, and thus joint spin measurements along different axes are undefined.

Since fields obey the principle of superposition, we can envisage superpositions of fields defined on different hemispheres. But fields defined on different hemispheres turn out to be *equivalent* to a field defined on a single hemisphere. Indeed it is easy to see that one can write for any axis u

$$F_{\Sigma_{+a}} \sim \cos\left(\frac{\theta_u - \theta_a}{2}\right) F_{\Sigma_{+u}} + \sin\left(\frac{\theta_u - \theta_a}{2}\right) F_{\Sigma_{-u}}, \quad (46)$$

meaning that although the two fields on the right and left hand sides of Eq. (46) are different – they are not defined on the same hemispherical surfaces –, they lead to exactly the same predictions. Indeed, when measurements are made along *any* axis b the averages of the left and right hand sides (hs) of Eq. (46) give the same result $\cos(\frac{\theta_a - \theta_b}{2})$. These fields thus define an *equivalence class*. From the particle standpoint, the field on the rhs of Eq. (46), F_{rhs} implies a different behavior: the no-perturbation axis is u , not a , and the particle distribution is not uniform. Hence there is a probability function $p_{F_{rhs}}(U = \pm \frac{1}{2}, \mathbf{r}) = 1$ or 0 depending on whether $\mathbf{r} \in \Sigma_{\pm u}$ and such that

$$P_{F_{rhs}}(U = \pm \frac{1}{2}) = \int p_{F_{rhs}}(U = \pm \frac{1}{2}, \mathbf{r}) \rho_{rhs}(\mathbf{r}) d\mathbf{r} = \cos^2\left(\frac{\theta_u - \theta_a}{2} + \frac{\pi}{4}(1 \pm 1)\right), \quad (47)$$

where $\rho_{rhs}(\mathbf{r})$ denotes the particle distribution when the field is given by the rhs of Eq. (46). However for $b \neq u$ there is no probability function $p_{F_{rhs}}(B = \pm \frac{1}{2}, \mathbf{r})$ hence $P_{F_{rhs}}(B = \pm \frac{1}{2})$ cannot depend on \mathbf{r} : the particle position does not ascribe probabilities and there is no sub-field mechanism that determines the outcome. This is consistent with Eqs. (44)-(45) in which the field rotation does not allow to define joint probabilities of the type $P_{F_{rhs}}(U = \pm 1 \cap B = \pm 1)$. Note that measuring A in the field F_{rhs} involves a perturbation in which the fields interfere due to the rotations in such a way as to obtain $P_{F_{rhs}}(A = -\frac{1}{2}) = 0$ irrespective of the initial the particle's position. The model cannot give a more specific interpretation in terms of the particle for this result produced by the interaction between the system and the apparatus measuring the spin projection along a .

Assume now an initial two-particle system is fragmented into two subsystems flying apart in opposite directions. Each of the two particles is embedded in a field defined on the surface of a small sphere. \mathbf{x}_1 (resp. \mathbf{x}_2) denotes the position of the subsystem 1 (resp. 2) sphere in the laboratory frame. The internal variables within each sphere are labeled by \mathbf{r}_1 and \mathbf{r}_2 . As soon as the fragmentation process is completed, the positions of each point-like particle as well as the fields are fixed, the spin of each system depending on the field distribution and the particle position on its spherical surface. The correlation of the particle positions must be set as $\mathbf{r}_1 = -\mathbf{r}_2$ in order to achieve $A_2 = -A_1$ for any axis a when there is no measurement perturbation. However the main element characterizing the correlations is the field distribution. For example the total field arising by correlating $F_{\Sigma_{+a}}^1$ defined on subsystem 1's sphere with Σ_{-a}^2 on subsystem 2 is given by $F_{\Sigma_{+a}}^1(\mathbf{r}_1) F_{\Sigma_{-a}}^2(\mathbf{r}_2)$. If in addition we also require the correlation $F_{\Sigma_{-a}}^1 \leftrightarrow F_{\Sigma_{+a}}^2$ the total field is given by the expression

$$F_{\mathbf{R}}(\mathbf{r}_1, \mathbf{r}_2) = F_{\Sigma_{+a}}^1(\mathbf{r}_1) F_{\Sigma_{-a}}^2(\mathbf{r}_2) - F_{\Sigma_{-a}}^1(\mathbf{r}_1) F_{\Sigma_{+a}}^2(\mathbf{r}_2). \quad (48)$$

The definition of $F_{\mathbf{R}}$ is mathematically non-separable over the individual subsystem spheres; this means that the field is defined as a whole, jointly over the two spheres. This is the only way to account for correlations between more than two hemispheres²; without further specifications, non-separability has nothing to do with non-locality (the field is set at the source in the intersection of the past light-cones of both system's space-time location). Recall that non-separable functions are not exceptional in classical physics, eg the classical action for multiparticle systems is non-separable, but that does not make particle classical mechanics non-local.

Let us now investigate measurements along arbitrary directions c for particle 1 and b for particle 2, and consider

$$P_{\mathbf{R}}(C_1 = 1, B_2 = 1) = \frac{1}{2} \left| \left\langle F_{\Sigma_{[a+c]}}^1 \right\rangle_{+c} \left\langle F_{\Sigma_{[-a+b]}}^2 \right\rangle_{+b} - \left\langle F_{\Sigma_{[-a+c]}}^1 \right\rangle_{+c} \left\langle F_{\Sigma_{[a+b]}}^2 \right\rangle_{+b} \right|^2 \quad (49)$$

where $N = 2$ is the probabilities normalization factor. $P_{\mathbf{R}}$ is a two-outcome probability and consequently depends on the correlated local averages of the both subsystems' fields rotated by the local interaction of each field with

² Actually it is possible to replace the non-separable field by a separable, factorizable one if the field is allowed to take complex values [16]; the separable field is then expanded as the sum of two non-separable fields, one of which never contributes to the averages, and the other being $F_{\mathbf{R}}$.

the measurement apparatus along the axes c and b . It can be simplified by using the expressions employed for the single particle averages: we then see that the expression between $|\dots|$ reduces to $\cos(\frac{\theta_b - \theta_c}{2})$ and is independent of a . Therefore $P_{\mathbf{r}}$ does not depend on the direction a of the single-particle fields that define $F_{\mathbf{r}}$ in Eq. (48). This implies the equivalence between fields $F_{\Sigma_{+a}}^1 F_{\Sigma_{-a}}^2 - F_{\Sigma_{-a}}^1 F_{\Sigma_{+a}}^2$ defined by different directions a , i.e. for any $b \neq a$

$$F_{\Sigma_{+a}}^1 F_{\Sigma_{-a}}^2 - F_{\Sigma_{-a}}^1 F_{\Sigma_{+a}}^2 \sim F_{\Sigma_{+b}}^1 F_{\Sigma_{-b}}^2 - F_{\Sigma_{-b}}^1 F_{\Sigma_{+b}}^2. \quad (50)$$

Both of these fields lead exactly to the same predictions for measurements along arbitrary axes and can thus not be distinguished. We will denote the left and right hand sides of Eq. (50) by $F_{\mathbf{r}(a)}$ and $F_{\mathbf{r}(b)}$ respectively. One consequence is that when computing $P_{\mathbf{r}}(A_1, B_2)$ one can use any of the two forms (50). Employing $F_{\mathbf{r}(a)}$ allows to make a conditional inference for B_2 , given that the measurement yielding A_1 does not perturb subsystem 1 and thus reveals to which hemisphere $\Sigma_{\pm a} \mathbf{r}_1$ belonged previous to the measurement:

$$P_{\mathbf{r}(a)}(A_1 = \frac{1}{2}, B_2 = \frac{1}{2}) = P(A_1 = \frac{1}{2})P(B_2 = \frac{1}{2} | A_1 = \frac{1}{2}) = P(\mathbf{r}_1 \in \Sigma_{+a}^1)P(B_2 = \frac{1}{2} | \mathbf{r}_1 \in \Sigma_{+a}^1) \quad (51)$$

$$= P(\mathbf{r}_1 \in \Sigma_{+a}^1)P(B_2 = \frac{1}{2} | \mathbf{r}_2 \in \Sigma_{-a}^2) = P(\mathbf{r}_1 \in \Sigma_{+a}^1)P_{F_{\Sigma_{-a}}^2}(B_2 = \frac{1}{2}). \quad (52)$$

The last step yields $P_{F_{\Sigma_{-a}}^2}(B_2 = \frac{1}{2})$ which is a single subsystem probability; this step is justified by the fact that given $\mathbf{r}_2 \in \Sigma_{-a}^2$, the field over subsystem 2 is equivalent, as discussed below Eq. (47) to $F_{\Sigma_{-a}}^2$. Eqs. (51)-(52) can be repeated by employing $F_{\mathbf{r}(b)}$, giving $P_{\mathbf{r}(b)}(A_1, B_2)$ in terms of a conditional probability inferred from a no-perturbation measurement along b (for subsystem 2). As in the single particle system case each particular realization of an equivalence class gives rise to different, incompatible, accounts grounded on the measurement that does not disturb the original field. Here however the equivalence class holds relative to the two-particle system, but relative to a single subsystem the specific form taken for $F_{\mathbf{r}}$ has different implications regarding the relation between the particle position and a given outcome. This does not affect a single subsystem probabilities, $P(A_i) = \frac{1}{2}$ for any a , but conditional probabilities can only be explicated when the form of the field corresponds to a measurement axis. For example $P_{\mathbf{r}(b)}(A_1 = \frac{1}{2} | C_2 = \frac{1}{2})$ cannot be computed: it is not correlated with the particle positions and due to the subsystems-apparata interactions, no inferences can be made; but $P_{\mathbf{r}(c)}(A_1 = \frac{1}{2} | C_2 = \frac{1}{2})$ can be inferred in terms of a single subsystem probability. The situation was the same for the single particle system described above, where no elementary probability could be ascribed to compute $P_{F_{\mathbf{r}(b)}}(B)$.

In this model, the particles' positions thus appear as pre-determined but only determine the outcome when there is no field perturbations from the apparatus interaction. The field configurations can also be taken as hidden variables and they do ascribe probabilities but only as members of an equivalence class that does not give a more complete specification than afforded by the quantum-mechanical state. The first implication is that there is no pre-existing outcome as an element of reality, even when it is possible to make a prediction with unit probability (in this case also there is an infinity of field/particle configurations giving that outcome). The second is that a given field/particle configuration (even if known) does not allow to specify sub-quantum probabilities for measurements along arbitrary axes. Any prediction that would complete quantum-mechanics, like the inference made on one subsystem's outcome once the other outcome is known, relies on a specific (but fictitious) field/particle configuration for which one of those measurements does not give rise to perturbations. This model therefore expels the locality issue from the EPR paradox (completeness of QM or simultaneous existence, based on locality, of physical quantities associated with non-commuting observables): by equating a quantum state with an equivalence class comprising an infinity of possible field-particle configurations the model denies the simultaneous existence of those physical quantities regardless of the locality issue, while keeping the physical predictions invariant. Note that the idea of a quantum state as being a label for an ensemble of underlying phenomena appears naturally when classical fields are considered as constituting a sub-quantum level of description [17, 18].

CONCLUSION

Bell-type hidden variable models give an explicit mathematical formulation of the EPR's simultaneous reality requirement concerning incompatible quantities. These models are irrelevant to quantum mechanics in so far as the latter denies the existence of joint distributions for incompatible quantities. We have argued that it is possible to uphold

non-commutativity and locality simultaneously and constructed to that effect two different types of models. However these models despite being local fail both Einstein's goal of efficiently *completing* quantum mechanics and Bell's goal of implementing explicitly *causality* to describe quantum correlations.

REFERENCES

1. J.S. Bell, *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, Cambridge, 2004), Chap. 4. [Original publication in *Foundations of Quantum Mechanics* New-York:Academic, 1971), p. 171-181].
2. J. F. Clauser and A. Shimony, Rep. Prog. Phys. 41, 1881 (1978).
3. P. Rastall, Phys. Lett. A 86, 85 (1981).
4. A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. 47 777 (1935).
5. A. Fine, J. Math. Phys. 23, 1306 (1982).
6. W. M. de Muynck and O. Abu-Zeid, Phys. Lett. A 100, 485 (1984).
7. L. Accardi, K. Imafuku and M. Regoli, arXiv:quant-ph/0112067 (2001) and Refs. therein.
8. L. J. Landau, Phys. Lett. A 120, 54 (1987).
9. M. Golshani and A. Fahmi, Annales de la Fondation Louis de Broglie, 26, 735 (2001).
10. E. P. Wigner, Am. J. Phys. 38, 1005 (1970).
11. G. C. Ghirardi and L. Marinatto, Phys. Lett. A 372, 198 (2008).
12. J.S. Bell, *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, Cambridge, 2004), Chap. 24 “La nouvelle cuisine”.
13. A. Matzkin, J. Phys. A 41 085303 (2008) [available as Eprint arXiv:quant-ph/0703251v3].
14. A. Matzkin, Phys. Rev. A 77 062110 (2008) [available as Eprint arXiv:0709.2114].
15. D. Aerts, Helv. Phys. Acta, 64, 1 (1991).
16. A. Matzkin arXiv:0808.2420v2 (2008).
17. A. Khrennikov J. Phys. A. 38 9051 (2005).
18. A. Khrennikov Phys. Lett. A 372 6588 (2008).